The Sweet Spot: How People Trade off Size and Definition on Mobile Devices

Hendrik O Knoche
University College London
Gower St
London, WC1E 6BT, UK
+44 207 679 3642
h.knoche@cs.ucl.ac.uk

M Angela Sasse
University College London
Gower St
London, WC1E 6BT, UK
+44 207 679 7212
a.sasse@cs.ucl.ac.uk

ABSTRACT
Mobile TV can deliver up-to-date content to users on the move. But it is currently unclear how to best adapt higher resolution TV content. In this paper, we describe a laboratory study with 35 participants who watched short clips of different content and shot types on a 200ppi PDA display at a resolution of either 120x90 or 168x128. Participants selected their preferred size and rated the acceptability of the visual experience. The preferred viewing ratio depended on the resolution and had to be at least 9.8*H*. The minimal angular resolution people required and which limited the up-scaling factor was 14 pixels per degree. Extreme long shots were best when depicted actors were at least 0.7° high. A second study researched the ecological validity of previous lab results by comparing them to results from the field. Image size yielded more value for users in the field than was apparent from lab results. In conclusion, current prediction models based on preferred viewing distances for TV and large displays do not predict viewing preferences on mobile devices. Our results will help to further the understanding of multimedia perception and service designers to deliver both economically viable and enjoyable experiences.

Categories and Subject Descriptors
H.5.1 Multimedia Information services – Evaluation/ methodology. H5.2 User Interfaces [user-centered design]

General Terms
Human Factors, Experimentation, Measurement, Design

Keywords
Mobile multimedia consumption, resolution, size, trade-off

1. INTRODUCTION
Advances in the development of displays have equipped mobile devices with 200ppi displays offering VGA (640x480) resolution. This capability goes some way towards reducing worries that coarse displays at short viewing distance will affect the experience of TV viewing. At a constant viewing distance (*D*) perceived video quality is determined by the size of the depicted image and its resolution. By increasing the viewing distance, the size decreases and the angular resolution of the picture increases. Viewing distances have been a central topic for TV video quality research and the introduction of HDTV. With mobile TV, however, the viewing distance can be considered fixed at about arm's length and the size of the picture and angular resolution are becoming key parameters. Previous research focused on identifying configurations resulting in maximum objective or subjective video quality, assuming that these regardless of e.g. size and immersion coincide with user preferences. However, recent lab-based research [1] has indicated that people are not basing their preferences for viewing distance on the best attainable subjective video quality when consuming TV content. But that the viewing ratio of distance to picture size needs to be considered. Furthermore, the observed viewing distances in the home [2] are even further away from preferences obtained in lab settings.

In this paper we describe two studies conducted to determine viewing preferences for mobile TV in two different settings. The first study investigated user preferences in trading off between image size and resolution for different content and shot types in a lab-based setting, specifically preferences for extreme long shots (XLS), which are important for content adaptation of sports content. A second study evaluated these parameters in a field trial, in which users viewed the content while traveling on underground trains. The results provide insights for 1) the development for displays, and 2) the optimal delivery of mobile TV content, especially how mobile TV content should be presented (up-sampled) on displays that are used at close distance on small mobile devices.

In the next section we present the background literature on viewing distances and resolution. The two studies are then described in Sections 3 and 4, followed by a discussion of the results in section 5. Conclusions are presented in Section 6.

2. BACKGROUND
When people who have not experienced mobile TV are asked about it, they usually mention the screen size as an obstacle to a satisfying experience [3], [4]. Talking about the size or the resolution of videos and displays only makes sense in conjunction with a viewing distance (*D*). When considering the viewing distance at which mobile devices are used, the relative size in terms of the viewing ratio (*VR*) – the quotient of *D* and screen height *H* – are not radically different from those in a home or PC TV setting. The real difference is the resolution of the content that is delivered to mobile devices, compared to standard television (SDTV). Historically, the viewing distance was the only way for people to adjust their preferred *VR* - the angular size of the

MM'08, October 26–31, 2008, Vancouver, British Columbia, Canada.

display - and the angular resolution, since most devices would only depict content at a fixed resolution. With advanced coding schemes and presentation devices that can stretch video, the question of the preferred angular resolution and viewing ratio can be rephrased. Many displays will allow for adjustment of the size of video content depending on user preferences.

2.1 Viewing at Close Distance

The amount of detail resolvable by the human eyes is primarily limited by the density of the light-sensitive rods and cones on the eyes' retina. Normal 20/20 vision is classified as the ability to resolve 1 minute of arc (1/60º) [5] and translates to 60 pixels per degree (ppd). The maximum amount of pixels p_{max} that can be resolved by a human at a given viewing distance D and a picture height H can be computed by the following equation:

$$p_{max} = H/D \times 2 \tan(1/120)$$

Viewing distance is often expressed in terms of the ratio between the distance of the observer to and the height of the visible screen. A viewing ratio of 5H describes a viewing distance five times the picture height (H). The visual angle (*VA*) or angular size (AS) θ expresses the viewing ratio in degrees regardless of D as illustrated in Figure 1.

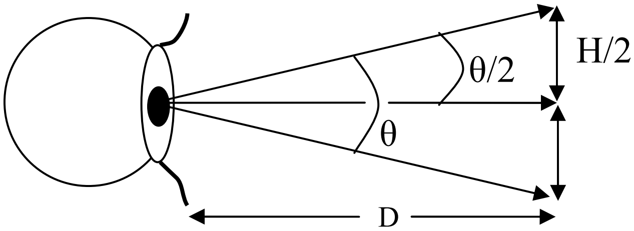

Figure 1: Viewing ratio and the visual angle

The human visual system uses two mechanisms to focus on objects: convergence and accommodation. Convergence denotes the eyes moving inward when focusing on nearby objects, and accommodation describes the focusing on objects of different distance by means of physically deforming the lens of the eye. The resting point of accommodation (RPA), i.e. the default distance at which objects appear sharp, e.g. when opening the eyes, is around 75cm for younger people and increases in distance with age [6]. The resting point of vergence (RPV) is 114cm when looking straight ahead, and drops to 89cm when looking 30 degrees down. This is a posture often seen in mobile TV consumption because people use their legs or bags on their laps as support for the hand holding the device (cf. Figure 2) [4]. The stress of convergence contributes more to visual discomfort than the stress of accommodation. Continued viewing at distances closer than the RPV can contribute to eyestrain [6]. When viewing distances come close to 15cm, people experience discomfort [7]. Boff & Lincoln showed that visual acuity decreases as viewing distance increases [8] so for close viewing distances people's acuity is at its maximum. Clearly, multimedia consumption on mobile devices happens at close range (cf. Figure 2), but its exact preferred viewing distance (PVD) has not been researched in relation to different sizes and resolutions.

In a laboratory study Knoche et al. found no evidence that people change their viewing distance in response to varying video sizes at constant angular resolution on a 115ppi mobile devices but kept an average viewing distance of about 40cm. Kato et al. obtained typical viewing distances of approx. 35cm or 11H from both standing and sitting people using a 166ppi mobile device [9].

Figure 2: Close range viewing of multimedia content

From research on traditional TV viewing we know that a number of factors can influence people's decision on their PVD. Early research by Thompson on preferred TV viewing distance suggested that people chose their viewing distance so that the TV lines were not visible anymore [10]. More recent research, however, has refuted this assumption. Due to the layout of the average living room, people typically watch TV at the so-called Lechner (US, 9ft) or Jackson (Europe, 3m) distance [11]. Unfortunately, both of these values are poorly documented and their original sources are not readily accessible. Nathan et al. showed that the viewing distance of regular TV in the home varied with the age of the viewers. The average viewing distance for 17 year olds and younger was 2.25m (7.8H), whereas adults watched from 3.37m (11.7H) [2]. The study did not explain this difference, but it did report that children were more mobile than adults, and much less likely to sit or lie on furniture while watching TV.

In a series of five studies Lund showed that participants' preferred viewing ratio was not a constant 7H. With increasing image size, and independent of resolution, the preferred viewing ratio approached 3H or 4H [1]. Based on Yuyama's [12] and his own results Lund hypothesized that viewers might select their viewing distance not to maximize perceived visual quality but "*to optimize a sense of presence or reality*" [1]. Ardito found that when brightness was reduced, there was a trend of participants sitting closer to the screen [13]. When watching HDTV content on a 38 inch screen (in a completely dark room), the average preferred viewing ratio was 3.8H, compared to 6.3H when viewing the same footage in brighter surroundings.

Ardito found the viewing distances for moving picture content to be further away than for still picture content. However, this effect and the contribution of the screen resolution were small in comparison to the large effect of the size of the screen on the PVD. For HDTV content, he predicted a viewing distance (in cm) of $D = (3.55H + 90)/H$. Although he did not test small mobile screens, he interpolated from a range of HDTV screen heights from 198cm to 15cm that for screens with a screen height close to zero the viewing distance would be 90cm [13].

2.2 The Effects of Size and Resolution

Since the early works of Kell et al. research has addressed TV resolution requirements in conjunction with size to improve the experience of the viewers [6]. Jesty found evidence for a preferred viewing distance [14] when participants chose their preferred viewing distance of projected pictures. Their preferred viewing ratio was constant for a given resolution. Ribchester argued that this could be merely attributed to conditioning to existing physical setups in the home [15]. Westerink & Roufs researched the subjective image quality of a range of angular resolutions (5, 17, 50 and 64ppd) by defocusing a projector lens and picture sizes 24, 48, 72 and 92cm at three viewing distances 2.9m, 3.9m and 5.4m – equivalent to viewing ratios between 22.5*H* and 3.2*H* - using projected stills in a dark room. They found that at a constant viewing distance the subjective quality of still pictures was influenced independently by both the resolution and the size of the pictures [16]. The subjective quality reached a maximum when the resolution equalled 16 cycles per degree (cpd) independent of the picture width. For pixel-based displays this translates to 32 ppd, which indicated that the gains in perceived visual quality from achieving a higher visual resolution beyond 16 cpd were not big enough to compensate for the detriment from the reduction in picture angle. Based on Westerink & Roufs' data, Barten inferred the square-root integral that predicted the effect of both picture size and picture resolution on perceived image quality [17]. However, the original data was based on still pictures and viewing distances between 2.9m and 5.4m and might not apply to consumption on rasterized mobile devices.

Sugama et al. found that for pictures of identical angular resolution of 27ppd – all shown at 6*H* - the subjective video quality on a 100ppi display was higher when they were viewed at a close distance of 40cm in comparison to viewing distances of 80cm and 1.6m [18]. However, the study did not control for the 'graininess' or pixilation of the display. At the closest distance the angular resolution of the display was 27ppd but for the largest viewing distance (1.6m) with the medium (54ppd) and large (100ppd) images the pixels were close to and above the human discrimination threshold. In a study that used different viewing distances (1.5, 2 and 3m) Hatada et al. showed that the angular size of the display was not sufficient to describe the effect of display size but that the absolute picture size or the absolute viewing distance needed to be considered [19]. Yu *et al.* found no statistical difference in assessors judging video quality impairments when the material was presented at 3*H* or 5*H* [20]. Lombard et al. found that bigger screens (115cm height) displaying standard definition television (SDTV) resulted in more intense experiences for the audience compared to smaller screens (30cm) but the level of enjoyment remained unchanged. Reeves et al. found that large HDTV presentations with high audio resulted in more positive evaluations in comparison to SDTV, e.g. in terms of excitement [21]. Apparently, larger depictions need to be coupled with higher resolutions in order to achieve higher levels of enjoyment.

Knoche et al. found that the perceived acceptability of clips' video quality played at their native resolution on a 115ppi mobile device decreased non-linearly when the dimensions decreased from 168x128 to 120x90 [22]. Either the clips' size of 20mm height (20*H*) or its resolution of 120x90 were too small for an acceptable experience of mobile TV. Size and resolution although correlated represent two different dimensions. Current models on video image quality e.g. [17] lack empirical results in the domain of mobile viewing devices. In summary almost all of the existing research on size and image resolution preferences for TV content was based on large screens and there is a gap in the understanding of what constitutes people's preferred angular resolution and viewing ratio when watching video on mobile devices.

2.3 Shot types

In this paper we use Thompson's classification [23] for Medium Close-Ups (MCU), Medium Shots (MS), Long Shots (LS), Very Long Shots (VLS) and Extreme Long Shots (XLS). Faced with the more constrained visual real estate, content producers are considering a different mix of shot types for mobile TV. In Asia, content creators produce soap operas especially for mobile devices, which are short and rely heavily on close-up shots with little dialogue. Most emotions have to be conveyed by means of facial expressions and "*there is very little dialogue and a lot of close-ups of characters striking exaggerated poses*" [24]. ESPN minimizes the use of extreme long shots In sports coverage for mobile devices [25]; instead it uses more high-lights with close-up shots. Little research has been done on how small screen sizes and resolutions might affect the presentation of shot types. In a first study by Knoche et al. this depended on the content type [26].

2.4 Lab vs. Field

Evaluation in the field is time-consuming and difficult to carry out because experimental control in terms of interruptions, movement, lighting and sound conditions is hard to achieve [27]. For video quality assessment the ITU guidelines suggest control of, e.g. lighting conditions [28], which is not feasible in the field. However, results that are obtained in more realistic settings have greater predictive validity because they are closer to the user's real experience [29]. The only previous research that compared video quality assessments in the field and the lab was a recent study by Jumisko-Pykköö & Hannuksela that evaluated the effect of packet loss on audio-visual quality [30]. In general participants rated the acceptability and of audio-visual quality higher in the field than in the laboratory for all four tested error ratios (from 1.7% to 20.6%). The same was true for the satisfaction ratings except for the 1.7% error ratio condition. Satisfaction ratings in the field (6.5) were lower than in the lab (7.5) while the participants' acceptability ratings for the same clips were higher for the field (89%) than in the lab (82%). The reason for this difference is not yet understood.

3. STUDY ON RESOLUTION AND SIZE

3.1 Material

From previous studies on multimedia consumption on mobile devices we obtained news [31], animation, music and sports video clips [22]. These base clips were based on footage from DVDs and terrestrial digital video broadcasts (DVB-T). For each of the four content types we used four shots of each of the shot types MS/MCU, LS, VLS and XLS (see Figure 3 for examples).

Our material did not provide consecutive shots of all types lasting for more than ten seconds, so in order to control for effects due to shot types each shot lasted for only 8-10 seconds. Due to differences in content, the most detailed shot types were not identical. For example, the football shot with the most detail was closer to a mid-shot than a medium close-up, and the most detailed shot in animation was closer to a close-up. Another difference worth pointing out is that the XLS of football and news depicted people far away whereas the XLS of music (moving camera) and animation (static shot) depicted a landscape.

Using previously tested material allowed for a comparison with earlier results [31] and [22], which partly addressed size and resolution concerns. The clips were originally encoded at 192kbps WMV V9 at a nominal 12.5 fps at two respective resolutions 120x90 and 168x126 with Audio V8 at 32kbps. The two sets of clips were then encoded at the six dimensions of 480x360, 400x300, 320x240, 240x180, 168x126 and 120x90 at a higher encoding bitrate in order to ensure that the resulting clips had the same visual quality. They appeared bigger on the screen but at the same resolution. The original pixels stretched over more pixels on the display. The text contained in the news clips was legible at all above sizes.

Figure 3: Shot type examples of animation, music football and news from left to right: MCU/MS, LS, XLS, VLS

The dimensions and values for angular size, angular resolution, and viewing ratio are summarized in Table 1. All values are based on 32cm viewing distance, which was the average observed in Study 1. AR and VR are rounded values. The encoding bitrate of 192kbps was chosen based on previous results [22] that showed that for 168x126 clips this resulted either in high acceptability ratings (over 70%) or that the ratings reached a plateau at this value.

Table 1: Experimental setup values (at D=32cm)

Video dim. in mm	Pixels used on display	*VR*	Ang Size	AR ppd 120x90	AR ppd 168x126
60 x 45	480 x 360	7*H*	7.4º	11	16
50 x 37.5	400 x 300	8.5*H*	6.1º	13	19
40 x 30	320 x 240	11*H*	4.9º	17	23
30 x 22.5	240 x 180	14*H*	3.7º	22	31
21 x 16	168 x 126	20*H*	2.6º	31	44
15 x 11.25	120 x 90	28*H*	1.8º	45	45‡

‡ The resolution of this footage was only 120x90 limited by the resolution of the display.

The clips were presented on an iPaq hx4700 with a 200 pixel per inch (ppi) 640x480 (VGA) resolution transflective TFT display with 64k colours. For content played at its native resolution and a viewing distance of 32cm this resulted in an angular resolution of 45ppd. The sound was delivered through a set of Sony MDR-Q66LW headphones. Each set of the six different size clips was arranged in a play list and played through the application The Core Pocket Media Player (TCPMP version 0.71). We checked that all clips played at their nominal frame rate using the benchmarking tool included in the TCPMP. Benchmarking videos encoded at 640x480 pixels showed that videos did not play at their nominal frame rate of 12.5 fps. The highest resolution that played at the nominal frame rate was 480x360, which was then chosen as the maximum for this study.

To better understand the effects of size and resolution for XLS on mobile devices we included four additional video clips. These had been produced in the context of a previous study that had looked at different zooms for content adaptation to mobile devices [32]. Two base XLS clips depicted football at two distances (the size of the depicted actors was different) and the other two clips were 1.6 times zoomed in versions of the base clips. The zoomed clips did not show all that was visible in the base clips but showed the content of a moving zoom window at a larger size. For further details on the preparation of the zoomed material consult [32].

This provided us with four different sizes of actors in the footage: 11, 15, 18 and 24 pixels in height in the original resolution 168x126. This would allow us to find out whether participants' preferences in terms of preferred size are due to the absolute size of the clips or depicted objects within the video clips. We prepared these four clips at the same six dimensions (see Table 1) but only one resolution 168x126 and at its original encoding bitrate - 350kbps WMV V9 at 12.5 fps and WMA V8 at 32kbps.

3.2 Procedure

The participants watched 16 randomized clips on a couch in a lab with ambient light of 345 lux. Each block of four clips had each content type and shot type appear at least once. The presentation assured that each content type and shot type combination was used at least once as the first clip. After the first 16 clips we showed the four XLS clips in randomized order which assured that the same base clips were not played twice in a row.

The instructions stated that the participants could assume any position sitting on a couch and that they deemed appropriate for following mobile TV. Each of the 20 clips (play lists) started playing at the smallest size. The participants were told to find their *favourite size* and point out which sizes they deemed *acceptable* and *unacceptable* in terms of the visual experience. They could use buttons to increase or decrease the size. On each button press, the video started over from the beginning. We encouraged and prompted the participants to explain why they found certain sizes unacceptable. Finding one's preferred size is based on the *method of adjustment* which was successfully adopted in previous video quality research by, e.g. Richardson et al. [33]. The method of acceptability was established by McCarthy et al. in [34].

We tested participants for visual acuity with a Snellen chart, and for colour-blindness with an Ishihara test. To capture participants' comments and measure viewing distances, participants were audio and video recorded. Viewing distance measures were also taken by means of a measuring stick that was occasionally held at the side of the participants, which did not seem to interfere with the participants' task.

3.3 Participants

A total of 35 paid participants (18f, 17m) with an average age of 25 took part in this study. Thirty participant had a visual acuity was 100% or better, 95% (1), 85% (1), 80% (1). Two male participants were colour-blind.

3.4 Results

For each video clip we obtained three measures - the *favourite size* at which participants preferred to watch, the *minimal size* and the *minimal angular resolution* (derived from the largest acceptable size) at which watching was still acceptable. We ran three mixed factor ANOVAs on *favourite size, minimal acceptable size* and *minimal angular resolution* as the dependent variables each with *content type* and *shot type* as *within*- and *resolution* as a *between-subjects* factor. The results are based on a total of 4200 acceptability and 700 favourite size ratings. The qualitative results are based on the 1030 comments we received.

Angular sizes are reported in degrees, resulting viewing ratios in terms of picture height (*H)* and angular resolutions in pixels per degree (ppd).

3.4.1 Viewing distance

Only one participant systematically varied the viewing distance with the six different size videos – pulling it closer for the smaller images. All other participants generally assumed the same posture when flicking through the different sizes. When they were unsure about the acceptability of a small size clip they occasionally pulled it closer for inspection but then usually changed back into their preferred position. We averaged the viewing distances of each participant during the trial. Both the average and the median of those average viewing distances were 32cm with a standard deviation σ of 6.8cm. Although the average viewing distance in the 168x126 resolution group was slightly higher (32.7cm; σ=6cm) than in the 120x90 group (31.8cm, σ =7.6cm) a t-test showed that this difference was not significant: t(33)=-0.372, n.s.

3.4.2 Acceptability of video quality

We averaged the acceptability scores of all participants for the six different sizes in both resolution groups (depicted in Figure 4).

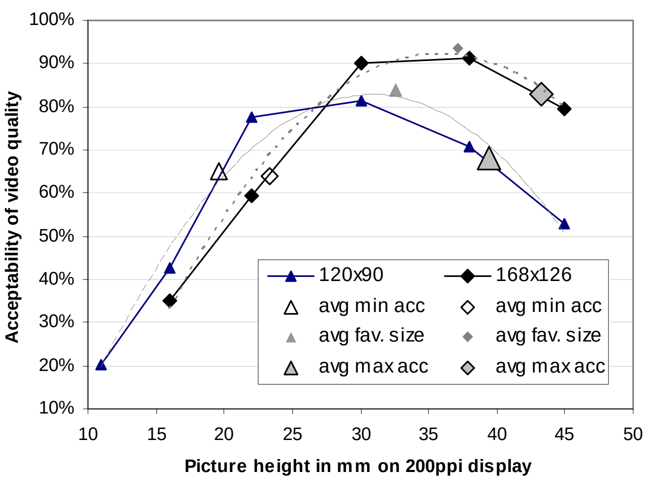

Figure 4: Video quality acceptability dependent picture height by resolution along with avg. min., favourite and max. size

The acceptability of the video quality varied tremendously with the size of the video. The averaged acceptability values of the video quality for both resolutions increased greatly for the larger sizes in comparison to the smallest size (picture height 11.25mm). However, the acceptability then reached a local maximum – 80% at 30mm picture height for the 120x90 resolution and 90% at 37.5mm picture height for the 168x126 resolution – after which the acceptability dropped off. The second order polynomial trend lines of the averaged acceptability scores were:

120x90: $y = -0.0016x^2 + 0.0988x - 0.6948$; ($R^2 = 0.985$),
168x126: $y = -0.0015x^2 + 0.1075x - 1.0051$; ($R^2 = 0.973$).

They result in local maxima of acceptability of video quality at a picture height of 31mm for 120x90 (10.3*H*, 16ppd) and 35.5mm for 168x126 (9*H*, 20ppd).

3.4.3 Favourite size

The participants in the higher *resolution* group had larger *favourite sizes* (F(1,33)=5.47, p<0.05). The average favourite sizes of all participants of the two resolution groups were 32.6mm (9.8*H*, 15ppd) and 37.2mm (8.6*H*, 19ppd) – slightly larger than the computed maxima of the polynomial trend lines in Figure 4 based on the averaged acceptability results. There was a significant main effect for *content type* F(3,99)=5.5, p<0.01. The Bonferroni adjusted pair comparisons showed that this effect was due to the news content with an average *favourite size* of 33mm (9.7*H*) - significantly smaller than football and music with an average *favourite size* of 35mm (9.1*H*). No significant effect was found for *shot type*. The interaction between *content type* and *shot type* was significant (F(9,297)=3.35, p<.01). Participants preferred to watch the XLS of football content at 39mm (8.2*H*) a significantly larger size compared to the XLS of animation and news at 35mm (9.1*H*).

3.4.4 Minimal size

We found a significant main effect for *shot type* (F(1,32)=40.71, p<0.001). The average minimal acceptable size of the more detailed shots LS was 19.5mm (16.4*H*) and 21mm for the MCU/MS (15.2*H*) significantly smaller than for XLS and VLS (both around 23mm, 13.9*H*). For the high *resolution* video clips at 168x126 the *minimal acceptable size* was 23.4mm – significantly larger than the 19.6mm (16.3*H*) for the low *resolution* clips (F(1,32)=7.32 p<0.05). In other words content at high *resolution* had to be presented at a larger size than at low *resolution* in order to be acceptable. There was a significant effect for *content type* (F(1,32)=7.32 p<0.05) on *minimal acceptable size.* The average *minimum acceptable size* for football content was 23mm (13.9*H*) but all other *content types* were still acceptable at 21mm (15.2*H*). An interaction effect between *shot type* and *resolution* (F(1,288)=10.78, p<0.001) (illustrated in Figure 5) showed that for the high *resolution* clips the differences in *minimal acceptable size* due to *shot type*s were more pronounced.

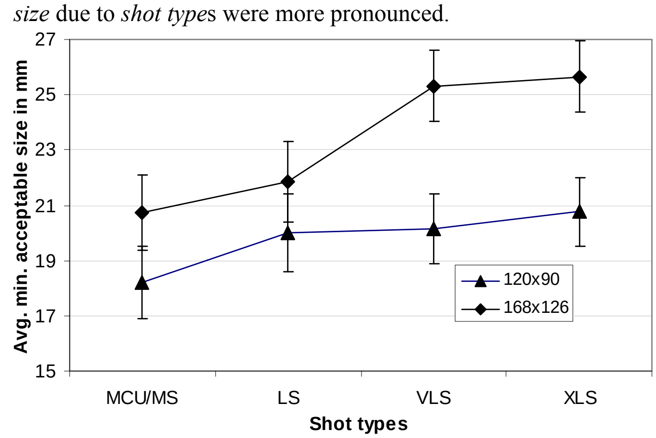

Figure 5: Interaction effect of *shot type* and *resolution* on *minimal acceptable size* of the picture

For the low *resolution* the only difference that remained significant was the required size for XLS (20.8mm) in comparison to the MCU/MS (18.2mm). Furthermore, an interaction between *shot type* and *content type* was based on individual clip differences - the animation's VLS, a relatively dark shot, the news' LS with the presenter being occasionally occluded and the football's XLS. They all required larger sizes to be acceptable. The animation's static LS was acceptable at smaller sizes than the other LS shots.

3.4.5 Minimal angular resolution

Resolution was the only factor that had a significant effect on the acceptable *minimal angular resolution* ($F(1,33)=7.05, p<.05$). The average lower bound was larger (17ppd) for the 168x126 group than for the 120x90 group (13.5ppd). We discuss the possibility of this being due to a ceiling effect in Sec 3.5.

3.4.6 Qualitative results

From the qualitative feedback we found that people deemed the smaller sizes unacceptable because they found them *"too small"*, *"couldn't figure out what's going on"*, "*hard to identify people*" and *"hard to look at"*. The number of these complaints (depicted in Figure 6, left) was marginal once the size reached 30mm in height (11*H*).

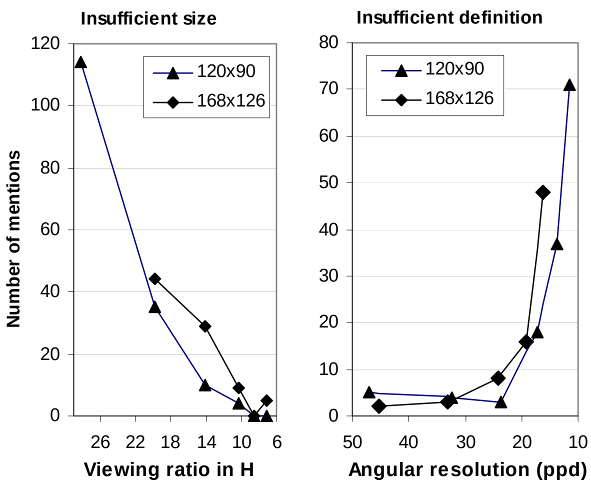

Figure 6: Participants' complaints about insufficient size (left) and insufficient definition (right)

Some participants commented that although the definition seemed high the size was not big enough to appreciate it. On the larger end the participants found the experience unacceptable because of the lack of definition or resolution once the angular resolution fell below 20ppd (see Figure 6, right). This problem was also made apparent by complaints about text albeit to a lesser degree. At small sizes (<22mm) participants complained about the effort required to read the text and at larger sizes, i.e. lower angular resolution (<17ppd), the quality of the text became too '*blurred*', '*pixelated*' or '*fuzzy*'. Other problems that people mentioned in favour of larger sizes were based on dark scenes, insufficient contrast, and movement either of the camera or in the scene. For all angular resolutions smaller than 24ppd more complaints came from the higher resolution group (which saw at a comparable angular resolution a larger picture than the lower resolution group).

3.4.7 Actor size in XLS

For the four clips that evaluated the influence due to *actor size* in XLS we followed the same approach in the analysis. We averaged the acceptability scores of all participants at all picture heights for the four clips to obtain the curves presented in Figure 7. XLS clips depicting actors that were larger in size were generally more acceptable at all sizes smaller than 37.5mm (8.5*H*) to be acceptable. The acceptability of all four clips reached its maximum at the two largest sizes. This means that the measures *favourite size* and *minimal angular resolution* are subject to possible ceiling effects. We ran *repeated measures one factor* ANOVAs on *favourite size*, *minimal* size and *minimal angular resolution* with *actor size* as the sole factor.

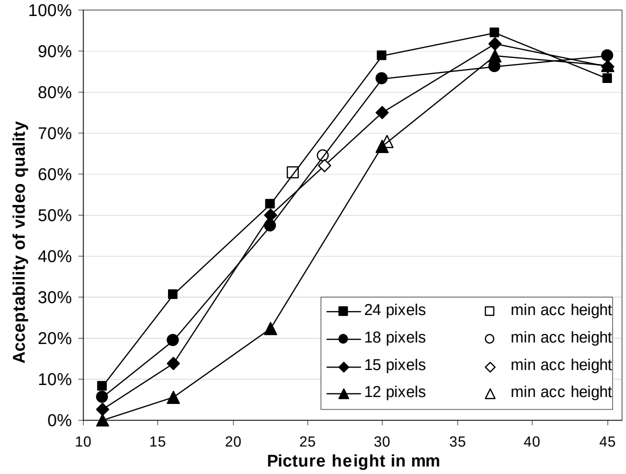

Figure 7: The acceptability of sports XLS clips by *picture height* for different *actor sizes*

As expected, clips with smaller *actor sizes* required higher *minimal acceptable sizes* than those with larger depictions – a significant linear effect ($F(3,102)=13.58$, $p<.001$). The angular size of the depicted actors for this lower bound ranged from 0.5º to 0.8º. The clips with smaller depictions of actors also had a higher *favourite size* than the clips with larger actor sizes ($F(3,102)=8.54$, $p<.001$). The mean *favourite size* increased from 38.5mm (18ppd) to 42.5mm (17ppd). The matching favourite angular sizes for actors were between 1.3º and 0.7º.

3.5 Discussion

3.5.1 Viewing distance

The viewing distances observed in this study are in line with earlier research by Kato et al. [9]. The fact that the viewing distance in this trial was smaller than in the previous study [22] that used the same content could be attributed to several factors.

1) The previously reported measures were obtained by estimating viewing distances based on observational video recordings. The measurements obtained in this study are more accurate.
2) Furthermore, in this study each clip started with a smaller image than in previous studies. This might have prompted many people to hold the device closer to their eyes. However, people were told that they should assume a comfortable posture that they would assume if they were watching mobile TV. The participants sat on a sofa, which might have affected their posture.
3) The higher resolution of the 200ppi display in comparison to the 115ppi display in the previous study.

3.5.2 Acceptability of video quality

In terms of *trading off* size and *definition* the acceptability of the video clips increased until their size translated to 10.6*H*, at which point the angular resolution was 16.5ppd for the 120x90, and about 8.7*H* for the 168x126 (19.4ppd) video clips. From there on, the acceptability declined and complaints about definition rose as angular size increased and angular resolution declined. Our participants made comments about the '*high definition*' at small size but did not try to achieve Westerink & Roufs' maximum video quality of 32ppd. Although angular resolution of 32ppd was possible to attain in both groups the resulting sizes were deemed too small. Apparently, size concerns must have been considered differently in our acceptability ratings since the computed acceptability maxima were close to the favourite sizes chosen by the participants.

The acceptability results showed that lower resolution content was more acceptable than higher resolution content at small sizes. This finding is consistent with Westerink and Roufs' [16]. In their study smaller depictions of low resolution still pictures received higher subjective quality marks than stills of larger size and higher resolution. In practice, this suggests that - if screens were not big enough - it would be counterproductive to deliver high-resolution content; using lower resolution would result in higher acceptability. This has tremendous implications for service providers, who could save on bandwidth and deliver a better experience to customers at the same time.

3.5.3 Minimal size

The different level of details portrayed in the studied shot types only made a difference when the lower limit at which the participants were willing to watch was reached. Shots that depicted their subject from closer up could be watched at smaller sizes. But again higher resolution called for larger sizes across all shot types. The participants accepted higher resolution displays of VLS and XLS only at larger sizes compared to the more detailed MCU/MS and LS. More research is required to explore the full extent of the interaction between resolution and shot types when minimal sizes are concerned.

3.5.4 Minimal angular resolution

The *minimal angular resolution* depended neither on *content* nor on *shot types*. The lowest acceptable angular resolution was around 14ppd - the same for all content and shot types. For up-scaling we do not have to consider these. This is close to the 11ppd that Lund found through the minimal viewing distances for projected video content in a dark room [1].

The effect of *resolution* – the minimal angular resolution for 168x126 was 17ppd and for 120x90 was 13.5ppd - could be due to a ceiling effect. The 168x126 group could not select larger sizes than were available and bring down the angular resolution more. The theoretical minimum at the largest size for 168x126 was 16ppd (11ppd for 120x90). The ceiling effect is also supported by the fact that the acceptability obtained from the polynomial trend line of the average maximal acceptable size (84%) depicted in Figure 4 is much larger than the values of all other bounds (between 63% and 71%).

3.5.5 Favourite size

The *favourite size* depended on the *resolution* of the content. Higher resolutions were preferably watched at larger sizes than lower resolutions. The average *favourite size* of *news* content (33mm) was smaller than of other content types. This could be rooted in perceived quality of text. People made the fewest complaints about text either being '*illegible*' or '*too hard to read*' at the 30mm picture size. In a previous study smaller depictions of news had received higher acceptability scores than larger depictions despite constant angular resolution [22].

Football on the other hand was preferred at significantly larger sizes due to the XLS and the MS. As explained the MS was less detailed than the comparable shots of the other content types, which were closer to an MCU. The XLS depicted a far away pitch in which actors were only 12 pixels in height in the original footage. At the *preferred size* the actors were about 0.7º tall.

The fact that we found significant interactions between content type and shot type could stem from the fact that possibly confounding factors were not controlled for. The qualitative feedback suggested that dark scenes, text, camera movement and the presence or absence of actors had an influence on the results.

In Figure 8, we have collated the preferred (PVD) and minimal viewing distances (VD) from the aforementioned studies by Lund, Ardito, Ardito et al. and Nathan et al. and plotted them in terms of the resulting viewing ratio and angular resolution. Results obtained in dark rooms are marked with shadows. The assumed lower limit of angular resolution is marked with a dotted black line. Our results were based on preferred viewing sizes (PVS) all others on preferred viewing distances (except for Lund-2 which included a minimal acceptable viewing distance).

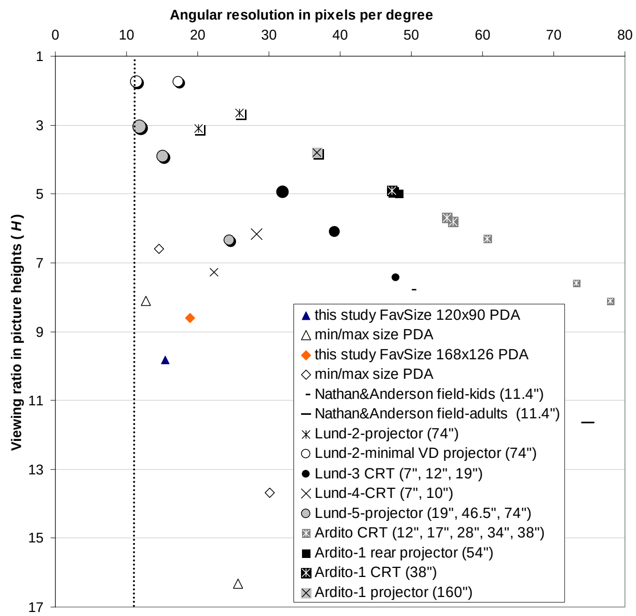

Figure 8: Comparison of results obtained by Lund, Ardito et al. and Nathan et al. on PVD with our results on preferred size

3.5.6 Actor size in XLS

In previous research [32] the preferred size of actors in XLS based on zooming into scenes was between 0.5º (176x144) and 0.7º (320x240). However, in that study increased actor sizes had to be traded off for a reduction in visual context by cropping the picture while the angular resolution of the picture was held constant (35ppd). In the study at hand increasing the size of players did not reduce context - only the angular resolution of the depicted video. The participants chose *angular sizes* of actors in XLS between 0.5º and 0.7º as the *minimal acceptable size*. But they favoured achieving larger *angular sizes* of actors between 0.7º and 1.4º and watched the footage at an angular resolution between 19ppd and 17ppd.

4. Study 2

In this study we evaluate the ecological validity of results obtained from two previous lab-based studies by showing the same clips on public transport.

4.1 Design

The experimental design followed the one used in [22] and [31]. We ran two groups: each group of 16 participants viewed 16 clips in groups of four, at each of the four sizes. The groups differed in whether they experienced *increasing* or *decreasing* image sizes. Within each group, we ran eight variations to control for content using a Latin squares design. This ensured that the different content clips were tested at each of the image sizes across participants.

4.2 Material

The video clips were encoded at four resolutions (240x180, 208x156, 168x126 and 120x90). Within each clip, the bitrate allocated to video was gracefully degraded every 20 seconds in steps of 32 kbps from a maximum of 224kbps down to 32kbps. The boundaries of these intervals were not pointed out to the participants. They were told only that the quality would vary over time and were presented with 16 clips, each of which gradually decreased in quality. For all clips the audio was encoded at 32kbps in stereo (WMV V9). The details about the productions of video clips can be found in [22] and for the news clips in [31].

4.3 Equipment

The test material was presented on an iPAQ 2210 with a 400Mhz X-scale processor, 64MB of RAM and a 512MB SD card. The screen was a 115ppi transflective TFT display with 64k colours and a resolution of 240x320. The iPAQ was equipped with a set of Sony MDR-Q66LW headphones to deliver the audio. We used the same interface as in the previous studies [22] and [31]. This interface offered two buttons that allowed the participants to switch back and forth between acceptable and unacceptable feedback with little effort.

4.4 Participants

Most of the 32 paid participants (11 women and 21 men, aged 20 to 65 with a median of 28 years) were university students. The majority came from the UK (20). English was the first language for 28 of the participants. Visual acuity was 100% or higher for 24, 95% for six, 90% for one and 85% for one of the participants.

4.5 Procedure

Before boarding the London Underground trains, participants were instructed by the experimenter, who accompanied them at all times. The participants were told that a technology consortium was investigating ways to deliver TV content to mobile devices, and that they wanted to find out the minimum acceptable video quality for watching news. The instructions stated: *"If you are watching the coverage and you find that the video quality becomes unacceptable at any time please click the button labelled 'Unacc'. When you continue watching the clips and you find that the quality has become acceptable again please click the button labelled 'Acc'... you can hold the PDA at any distance that is comfortable for you."* Each clip started with the interface in the 'Acc.' state.

The participants watched eight clips on the outbound journey, and another eight clips on the return train. The train journeys included both underground and over-ground segments. Throughout the experiment on the trains the participants were video recorded and a debrief interview concluded the session.

4.6 Results

We combined the data obtained in this experiment with data from two previous lab studies – from 64 participants from [22] and 32 participants [31]. Only results of those clips that were shown both in the lab and the field were included. The results were analyzed based on each 20 second encoding bitrate segment. If a participant judged the quality unacceptable at any time during a segment it was conservatively classified as unacceptable. We used a binary logistic regression to test for main effects and interactions between the independent variables of the previous studies – *Image Size*, *Video Bitrate* and *Content Type* and *Context*. Context denoted whether the data was obtained in the lab or the field. Control variables *Gender, isNativeSpeaker* and *Size Order* were included in the analysis.

The regression revealed significant effects on all of the control variables. Women found the video quality more acceptable than men [$\chi^2(1)=17.1$, $P<.001$], non-native speaker more than native speakers [$\chi^2(1)=8.6$, $P<0.001$] and the people whose clips increased in size more than those whose clips decreased in size during the experiment [$\chi^2(1)=119.9$, $p<.001$]. As in the previous studies *Image Size* [$\chi^2(1)=221.1$, $p<0.001$], *Video Bitrate* [$\chi^2(1)=16.7$, $p<0.001$], *Content Type* [$\chi^2(3)=1027.9.$, $p<.001$] were significant predictors of *acceptability*. Larger *Image Sizes* and higher *Video Bitrates* resulted in higher *acceptability*. But at lowest *Video Bitrate* the benefits of larger *Image Sizes* diminished. It turned out that *Context* was a significant predictor of *acceptability* [$\chi^2(1)=20.6$, $p<.001$] – the participants found the quality of the clips more acceptable on the trains than in the lab. This was not true across the board as the interaction of *Context* with *Image Size* was also significant predictor [$\chi^2(1)=16.4$, $p<0.001$] of *acceptability*. For the smaller *Image Sizes* there was no significant difference between the lab and the field but a non-parametric Kruskal-Wallis test [$\chi^2(1)=24.56$, $p<.001$] showed that the participants in the field found the larger two sizes more acceptable than the participants in the lab. This finding is summarised in Figure 9.

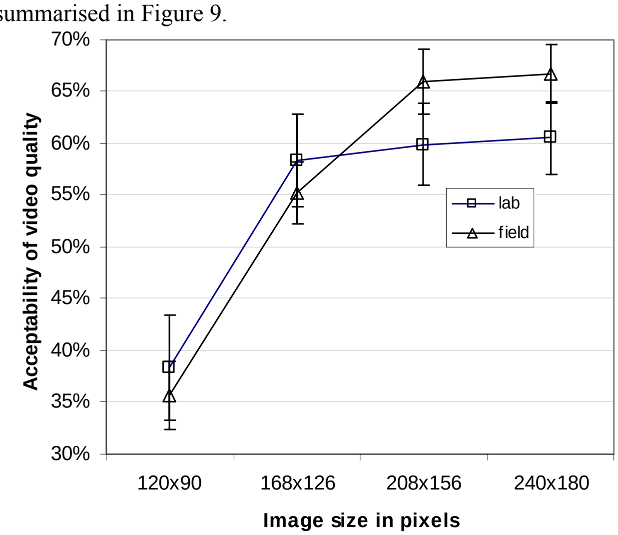

Figure 9: The interaction of *image size* and *context*

The interaction of *Context* with *Video Bitrate* was another significant predictor [$\chi^2(1)=20.2$, $p<.001$] of *acceptability*. At high *Video Bitrates* there was no difference between lab and field but for low *Video Bitrates* (<100kbps) the participants in the field found the video quality more acceptable than the participants in the lab. Figure 10 summarizes this interaction.

4.7 Discussion

We found that lab results on the acceptability of video quality were generally more conservative than their counterparts obtained in the lab. This is in line with the results from Jumisko-Pyykkö et al., in which participants rated the audio-visual quality of clips impaired by loss consistently higher in the field compared to the lab. The difference was most pronounced at the lowest quality – the highest loss ratio [30]. We found the same to be the case for low encoding bitrates. In terms of the size requirements the story was different. Our results showed that in the field the larger sizes (208x156 and 240x180) yielded a higher acceptability than in the lab. The acceptability of depictions smaller than 34.5mm resulted in equally reduced experiences both in the lab and the field. Further research is required to find the reason behind this.

From the view point of service providers delivering content in medium to high video quality, however, this should be no concern. The lab results should provide them with good conservative estimates that still hold in the field.

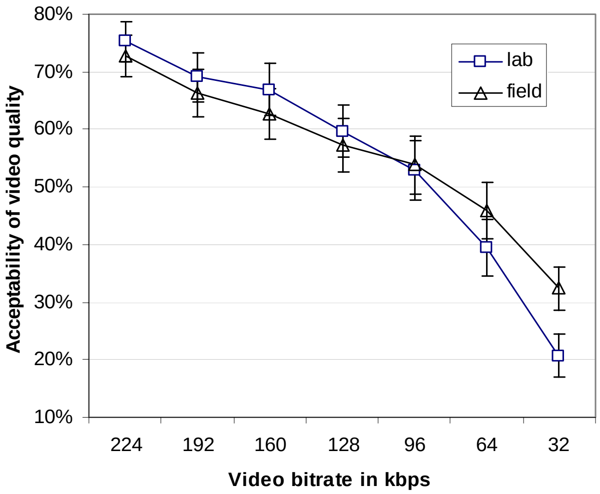

Figure 10: The interaction of *video bitrate* and *context*

5. OVERALL DISCUSSION

In order to compare the results from Study 1 with the field results from Study 2 and the lab results from [22] and [31], we weighted the acceptability scores of the different shot types from Study 1 according to the percentage with which they occurred in the footage in [22] and [31]. Figure 11 collates the results from Study 1 and 2 of this paper with the previous lab results [22] and [31].

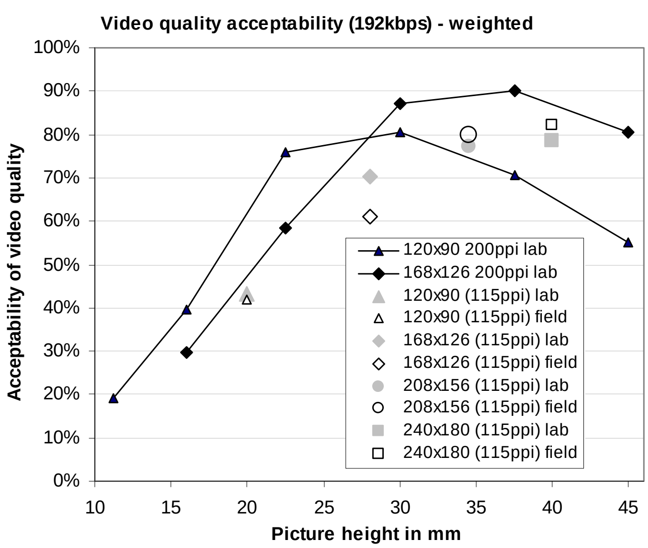

Figure 11: Video acceptability of Study 1, 2 and the combined lab results of [22] and [31] (in grey) by picture height

The acceptability results of 120x90 and 168x126 clips in the previous lab studies on a 115ppi device were lower than those obtained in this study (200ppi) but follow the same trend. The discrepancy could stem from the difference in displays and experimental procedures. The acceptability ratings of the clips at 208x156 and 240x180 from the previous lab study are below those of the 168x126 clips of this study in line with what the results in Sec. 3.4.2 would make us expect. If these higher resolution clips were increased in size they should surpass the acceptability ratings of 168x126 assuming that the bitrate of 192kbps can sufficiently encode the spatial information. This is a general limitation to our results on angular resolution requirements. We cannot know the exact resolution of the content due to the spatio-temporal compression of 192kbps.

6. CONCLUSION

Mobile TV services should be designed for close *viewing distances* between 25cm to 50cm. A distance of 32cm was the average in our study. Like [22] we found no adjustment of viewing distance depending on the resolution or the size of the footage. Mobile TV viewing distances might depend more on the posture of people within a given environment.

Both size and the available resolution of the content have to be taken into account for the best presentation of mobile TV material. Our results showed that participants' preferences for watching low resolution content depended first on size – they required at the very minimum sizes of 19.6mm (16.3*H*) for 120x90 resolution content but preferred sizes between 32.6mm (9.8*H*) and 37mm (8.6*H*) for 120x90 and 168x126 resolution content respectively. These values resulted in angular resolution of 15ppd and 19ppd. High resolution content had to be presented at sufficient sizes or resulted in a poorer overall experience than lower resolution content at the same size. Sizes that yielded the 32ppd – identified as optimal picture quality in [16] - did not coincide with the participants' favourite sizes but were criticized for being too small. The sizes of the computed acceptability ratings maxima, however, were close to the participants' favourite sizes. From our results it seems plausible that the preferred sizes of content with higher resolution than tested in our experiments will be preferred at even larger sizes. For devices with a screen size of 4cm height QCIF resolution content should result in the highest acceptability at a comparable encoding quality used in our study. A general limit for up-scaling video clips regardless of content and shot types was a resulting angular resolution of about 14ppd close to the 11ppd observed by Lund, derived from minimum viewing distances of large projections of TV content in a dark room [1].

Apart from the XLS, *shot types* were only a concern at the lower limits of acceptable size. MCU and MS could still be presented at smaller sizes than other shot types but their *favourite sizes* did not differ from other shot types. To rely on them in production would only make sense for content that would be shown on displays smaller than 22mm – the field results from Study 2, however, indicate that this will not be met by much user approval. The acceptability gains for XLS by zooming are substantial. Content adaptation employing zooming approaches should consider a resulting angular size of actors in XLS of at least 0.5° as a lower limit but ideally 0.7º and larger as long as the loss of context is within the limits as outlined in [32].

Although our research showed that lab experiments may be a conservative estimate of acceptability of video quality in the field this was not true for all observed factors. Especially for effects that are not fully understood yet – as in our case image size - tests in the field are advisable to validate, possibly correct and enhance laboratory results.

ACKNOWLEDGMENTS

This work was supported by the IST-2005-27034 UNIversal satellite home Connection (UNIC) project.

REFERENCES

[1] Lund, A. M. *The Influence of Video Image Size and Resolution on Viewing-Distance Preference* SMPTE, 102, p. 406-15, May 1993.

[2] Nathan, J. G., Anderson, D. R., Field, D. E., Collins, P. *Television viewing at home: Distances and visual angles of children and adults* Human Factors, 27, p. 467-76, 1985.

[3] Knoche, H. & McCarthy, J. *Mobile Users' Needs and Expectations of Future Multimedia Services* in Proceedings of the WWRF12, 2004

[4] Chipchase, J., Yanqing, C., Jung, Y. *Personal Television: A Qualitative Study of Mobile TV Users in South Korea*, 2007.

[5] Luther, A. C. *Principles of Digital Audio and Video* Boston, London: Artech House Publishers, 1996

[6] Owens, D. A. & Wolfe-Kelly, K. *Near Work, Visual Fatigue, and Variations of Oculomotor Tonus* Investigative Ophthalmology and Visual Science, 28, p. 743-9, 1987.

[7] Ankrum, D. R. *Viewing Distance at Computer Workstations* Work Place Ergonomics, p. 10-2, Sep/Oct 1996.

[8] Boff, K. R. & Lincoln, J. E. *Visual Acuity: Effect of Viewing Distance and Luminance Level,* In Engineering Data Compendium, Human Perception and Performance (Ohio: AAMRL: Wright-Patterson AFB, 1988

[9] Kato, S., Boon, C. S., Fujibayashi, A., Hangai, S., Hamamoto, T. *Perceptual Quality of Motion of Video Sequences on Mobile Terminals* in Proceedings of the IASTED International Conference, p. 442-7, 2005

[10] Thompson, F. T. *Television line structure suppression* SMPTE, 66, p. 603-6, 1957.

[11] Diamant, L. *The Broadcast Communications Dictionary* (3rd ed.) Greenwood Press, 1989

[12] Yuyama, I. *Fundamental Requirements for High-Definition Television,* In High-Definition Television NHK Technical Monograph (NHK Advanced broadcasting systems research division, 2008

[13] Ardito, M. *Studies on the Influence of Display Size and Picture Brightness on the Preferred Viewing Distance for HDTV* SMPTE, 103 (8) 1994.

[14] Jesty, L. C. *The relation between picture size, viewing distance and picture quality* in Proc.Inst.Electr.Eng., p. 425-39, 1958

[15] Ribchester, E. *Discussion Before th Radio and Telecommunication Section, 19th February 1958* Proc.Institution of Electrical Engineering, 105B, p. 437, 1958.

[16] Westerink, J. H. & Roufs, J. A. *Subjective Image Quality as a Function of Viewing Distance, Resolution, and Picture Size* SMPTE Journal, 98, p. 113-9, Feb1989.

[17] Barten, P. G. J. *Effect of Picture Size and Definition on Perceived Image Quality* IEEE Trans.on Electron Devices, 36 (9), p. 1865-9, Sept.1989.

[18] Sugama, Y., Yoshida, T., Hamamoto, T., Hangai, S., Seng, B. C., Kato, S. *A Comparison of Subjective Picture Quality with Objective Measure Using Subjective Spatial Frequency* in Proc.of ICME, p. 1262-5, 2005

[19] Hatada, T., Sakata, H., Kusaka, H. *Psychophysical analysis of the 'sensation of reality' induced by a visual wide-field display* SMPTE, 89, p. 560-9, 1980.

[20] Yu, Z., Wu, H. R., Ferguson, T. *The Influence of Viewing Distance on Subjective Impairment Assessment* IEEE Transactions on Broadcasting, 48 (4), p. 331-6, December2002.

[21] Reeves, B., Detenber, B. H., Steuer, J. *New Televisions: The Effects of Big Pictures and Big Sound On Viewer Responses to the Screen* in , 1993

[22] Knoche, H., McCarthy, J., Sasse, M. A. *Can Small Be Beautiful? Assessing Image Resolution Requirements for Mobile TV* in Proceedings of ACM Multimedia 2005, p. 829-38, ACM, 2005

[23] Thompson, R. *Grammar of the shot* Elsevier Focal Press, 1998

[24] Guardian, *Romantic drama in China soap opera only for mobile phones*., Guardian Newspapers Limited. at: http://www.buzzle.com/editorials/6-28-2005-72274.asp 2005

[25] Gwinn, E. & Hughlett, M., *Mobile TV for your cell phone*., Chicago Tribune. at: http://home.hamptonroads.com/stories/story.cfm?story=93423&ran=38197 2005

[26] Knoche, H., McCarthy, J., Sasse, M. A. *A close-up on Mobile TV: The effect of low resolutions on shot types* in Proc.of EuroITV '06 - Beyond Usability, Broadcast and TV, p. 359-67, 2006

[27] Tamminen, S., Oulasvirta, A., Toiskallio, K., Kankainen, A. *Understanding mobile contexts.* Special Issue of Journal of Personal and Ubiquitous Computing, 8 (135), p. 143, 2004.

[28] ITU-R *Methodology for the subjective assessment of the quality of television pictures.* (Rep. No. BT.500-7) , 2004.

[29] Sasse, M. A. & Knoche, H. *Quality in Context - an ecological approach to assessing QoS for mobile TV* in Proceedings of 2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of Systems, 2006

[30] Jumisko-Pyykkö, S. & Hannuksela, M. *Does Context Matter in Quality Evaluation of Mobile Television* in Proc.of mobile HCI, 2008

[31] Knoche, H., McCarthy, J., Sasse, M. A. *Reading the Fine Print: The Effect of Text Legibility on Perceived Video Quality in Mobile TV.* in Proceedings of ACM Multimedia, p. 727-30, 2006

[32] Knoche, H., Papaleo, M., Sasse, M. A., Vanelli-Coralli, A. *The Kindest Cut: Enhancing the User Experience of Mobile TV through Adequate Zooming* in Proceedings of ACM Multimedia 2007, p. 87-96, ACM, 2007

[33] Richardson, I. E. G. & Kannangara, C. S. *Fast subjective video quality measurement with user feedback* Electronics Letters, 40 (13), p. 799-801, 2004.

[34] McCarthy, J., Sasse, M. A., Miras, D. *Sharp or smooth? Comparing the effects of quantization vs. frame rate for streamed video* in Proc.CHI, p. 535-42, 2004